Leaf traits associated with drought adaptation in faba bean (*Vicia faba* L.)

Doctoral thesis

Hamid Khazaei

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Agriculture and Forestry, University of Helsinki, for public examination in lecture room B4 at Viikki campus, Latokartanonkaari 7–9, Helsinki, on 26 September 2014, at 12 o’clock noon.

Helsinki 2014
Supervisors:
Adjunct Professor Frederick L. Stoddard
Department of Agricultural Sciences
University of Helsinki
Finland

Professor Mikko J. Sillanpää
Department of Mathematical Sciences
Department of Biology
Biocenter Oulu
University of Oulu
Finland

Reviewers:
Assistant Professor Petr Smýkal
Department of Botany
Palacký University in Olomouc
Czech Republic

Adjunct Professor Simo Hovinen

Opponent:
Professor Wolfgang Link
Department of Crop Sciences
Georg-August-University
Germany

Cover figure: Faba bean germplasm grown in the greenhouse. © Hamid Khazaei

ISBN 978-952-10-8905-3 (Print)
ISBN 978-952-10-8906-0 (Online)
ISSN 1798-7407 (Print)
ISSN 1798-744X (Online)
ISSN-L 1798-7407
Department of Agricultural Sciences | Publications | 38
Electronic publication at http://ethesis.helsinki.fi

Unigrafia
Helsinki 2014
List of original publications

This thesis is based on the following publications, which are referred in the text with their roman numerals.


II. Khazaei H, Street K, Santanen A, Bari A, Stoddard FL (2013) Do faba bean (Vicia faba L.) accessions from environments with contrasting seasonal moisture availabilities differ in stomatal characteristics and related traits? Genetic Resources and Crop Evolution 60:2343–2357


Statement of contributions on these articles:

<table>
<thead>
<tr>
<th>Author Contributions</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceived and designed the experiments</td>
<td>HK KS FLS</td>
<td>HK FLS</td>
<td>HK FLS DOS</td>
<td>HK FLS</td>
</tr>
<tr>
<td>Performed the experiments</td>
<td>HK</td>
<td>HK</td>
<td>HK</td>
<td>HK</td>
</tr>
<tr>
<td>Analysed the data</td>
<td>HK AB</td>
<td>HK</td>
<td>HK DOS</td>
<td>HK DOS MJS</td>
</tr>
<tr>
<td>Contributed reagents/materials/analysis tools</td>
<td>HK KS AB</td>
<td>HK KS AB</td>
<td>HK DOS MJS</td>
<td>HK DOS MJS</td>
</tr>
<tr>
<td>Wrote the paper</td>
<td>HK KS AB</td>
<td>HK KS AS</td>
<td>HK FLS MJS</td>
<td>HK FLS MJS</td>
</tr>
</tbody>
</table>

HK Hamid Khazaei, KS Kenneth Street, FLS Frederick L. Stoddard, AB Abdallah Bari, MM Michael Mackay, AS Arja Santanen, DOS Donal M. O’Sullivan, MJS Mikko J. Sillanpää

The original publications (II, III and IV) are reprinted with the permission of Springer.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFLP</td>
<td>Amplified fragment length polymorphism</td>
</tr>
<tr>
<td>BC</td>
<td>Backcross</td>
</tr>
<tr>
<td>BSA</td>
<td>Bulked sergeant analysis</td>
</tr>
<tr>
<td>CGIAR</td>
<td>Consultative Group on International Agricultural Research</td>
</tr>
<tr>
<td>cM</td>
<td>CentiMorgans</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed sequence tag</td>
</tr>
<tr>
<td>EUW</td>
<td>Efficient use of water</td>
</tr>
<tr>
<td>FIGS</td>
<td>Focused identification of germplasm strategy</td>
</tr>
<tr>
<td>g</td>
<td>Stomatal conductance</td>
</tr>
<tr>
<td>ICARDA</td>
<td>International Centre for Agricultural Research in the Dry Areas</td>
</tr>
<tr>
<td>ISSR</td>
<td>Inter-simple sequence repeats</td>
</tr>
<tr>
<td>ITAP</td>
<td>Intron-targeted amplified polymorphism</td>
</tr>
<tr>
<td>KASP</td>
<td>Kompetitive allele specific PCR</td>
</tr>
<tr>
<td>LG</td>
<td>Linkage group</td>
</tr>
<tr>
<td>LOD</td>
<td>Logarithm (base 10) of odds</td>
</tr>
<tr>
<td>MAS</td>
<td>Marker assisted selection</td>
</tr>
<tr>
<td>PGR</td>
<td>Plant genetic resources</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative trait loci</td>
</tr>
<tr>
<td>RAPD</td>
<td>Randomly amplified polymorphic DNA</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RGA</td>
<td>Resistance gene analogue</td>
</tr>
<tr>
<td>RIL</td>
<td>Recombinant inbred line</td>
</tr>
<tr>
<td>SCAR</td>
<td>Sequence characterized amplified regions</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SSR</td>
<td>Simple sequence repeats</td>
</tr>
<tr>
<td>STS</td>
<td>Sequence tagged sites</td>
</tr>
<tr>
<td>TRAP</td>
<td>Target region amplification polymorphism</td>
</tr>
<tr>
<td>WUE</td>
<td>Water use efficiency</td>
</tr>
</tbody>
</table>
ABSTRACT

The potential for use of faba bean (*Vicia faba* L.) is increasing worldwide due to its positive environmental impact and nutritional interest, but there are many challenges for faba bean breeding and cultivation. These include its mixed breeding system, its unknown origin and wild progenitor, its large genome being the biggest diploid genome among field crops, and its relative sensitivity to biotic and abiotic stresses (e.g., drought). Little is known about the ecological adaptation of faba bean germplasm, or about the locations and effects of genes that influence traits related to drought adaptation, especially stomatal morphology and function as key characters for gas exchange between plant and atmosphere. Thus, the current study had two main goals, a) to examine whether faba bean germplasm from drought-prone and drought-free environments differed in leaf traits related to drought adaptation while testing a novel genetic resources utilization tool, and b) to screen the genome for regions and candidate genes controlling morpho-physiological traits related to drought adaptation.

Two sets of faba bean germplasm each containing 201 accessions from dry and wet regions of the world were chosen according to the principles of FIGS (the Focused Identification of Germplasm Strategy). Leaf morpho-physiological traits related to drought response (e.g., stomatal characteristics and water status) were measured under well watered conditions in a climate-controlled greenhouse. Thereafter, two subsets of 10 accessions each were exposed to water deficit. The significance of the difference between the two sets indicated the potential of FIGS to search for traits related to abiotic stress adaptation. Machine-learning algorithms and multivariate statistics showed that the discrimination between the two sets could be based on pertinent physiological traits, particularly leaf temperature and relative water content. When the plants exposed to water deficit, leaf temperature was the most responsive trait.

Four bi-parental populations were developed, of which Mélodie/2 × ILB938/2 showed the highest number of polymorphic single nucleotide polymorphisms (SNPs) and was advanced to the F₃ generation, in which 211 individuals were tested for 222 SNPs derived from *Medicago truncatula* sequence information. The population was phenotyped for several morpho-physiological traits during 2013 and 2014. In total, 188 polymorphic SNPs were assigned to nine linkage groups that covered ~ 928 cM with an average inter-marker distance of 5.8 cM. The map showed a high degree of synteny with the genome of *M. truncatula*. Most of the detected QTLs for stomatal morphology and function were in a single region of faba bean chromosome II syntenic with a segment of *M. truncatula* chromosome IV that harbours receptor-like protein kinase. Furthermore, a novel locus, *ssp1*, for stipule pigmentation was mapped in a well conserved region of *M. truncatula* chromosome V containing some candidate Myb and bHLH transcription factor genes.

The difference between the leaf temperatures of the two FIGS sets (the wet and dry set) allowed us to find a reliable and cost-efficient phenotyping tool for screening drought adaptation related traits in this species. Furthermore, using an appropriate mapping population and using novel *M. truncatula*-derived SNPs all brought success to detect the genetic regions and to indicate candidate genes. Furthermore, our results confirm the genomic data from model plant species can easily be translated to faba bean. Finally, breeding faba bean for drought adaptation can be made more straightforward by combining the use of germplasm tools such as FIGS, rapidly assessed phenotypes such as canopy temperature, and genomic tools for detecting candidate genes.
1 INTRODUCTION

1.1 Faba bean (Vicia faba L.)

Faba bean (Figure 1) also known in English as broad bean, field bean, fava bean or horse bean is an annual herbaceous species. It is belong to the family Fabaceae, which is the third largest angiosperm family, including 19,400 species and 740 genera (Lewis et al. 2005; Kew 2014). Faba bean is a member of subfamily Faboideae, tribe Vicieae and genus Vicia L. This genus is unusual in having a huge variation in genome size, from ~ 1,880 Mbp in V. monantha Retz. to 13,000 Mbp in V. faba (Raina and Rees 1983).

Figure 1 Faba bean. 1 flower, 2 stigma, 3 pod, 4 seed, 5 and 6 seed profiles, 7 pistil. Source: Thomé (1885)

Faba bean is on the first domesticated food legumes, probably in the early Neolithic period (8,000 B.C.). The word faba is originally from one of the roots of Greek verb, φαγέω, 'to eat', indicating its basic role as food in ancient Greek and Rome (Muratova 1931). The wild progenitor and place of origin of faba bean are still unknown (Cubero 1974; Maxted 1993), and no successful cross between faba bean and other Vicia species has been reported. A Near or Middle East centre of origin has been proposed for the faba bean (Cubero 1974), with radiation in four directions from the centre: Europe, along the North Africa coast to Spain, along the Nile Valley to Ethiopia, and from Mesopotamia to India. In addition, Afghanistan and Ethiopia have been proposed as secondary centres of diversity (Lawes et al. 1983), and subsequently China as well, since its faba bean gene pool is isolated from those of Ethiopia and western Asia (Zong et al. 2009). Faba bean is classified into four groups
according to seed size: minor (tick bean, ellipsoidal seed; 0.4 – 0.6 g / seed), equina (horse bean, field bean, flattened seed; 0.6 – 1.0 g / seed), major (broad bean, flattened seed; 1.0 – 3.0 g / seed) and paucijuga (a primitive form likely close to wild progenitor; 0.3 – 0.4 g / seed) (Muratova 1931; Hanelt 1972; Cubero 1974). The major type apparently emerged separately in the Mediterranean basin and China, the equina type mostly in North Africa and Middle East, and the minor and paucijuga types mostly in Ethiopia, northern Europe, Nepal and Bangladesh (Duc 1997).

In 2012, world production was 4.2 million tons (Tg) from nearly 2.5 million ha (M ha) (FAOSTAT 2014). China was the major producer, with ~1 M ha (38% of the world planted area) and 1.4 Tg (33% of world production), and Ethiopia was second (0.57 M ha). In the European Union, its cultivated area was 0.34 M ha with production of 0.97 Tg (Eurostat 2014). Faba bean is the fourth most important cool-season food legume in the world after chickpea (Cicer arietinum L.), pea (Pisum sativum L.) and lentil (Lens culinaris Medik.) (11.6, 9.8 and 4.5 Tg, respectively), but all of these are produced on a much smaller scale than soybean (Glycine max [L.] Merr.) (241.8 Tg). Recently, interest in growing faba bean has greatly increased in Europe, Australia and North America because of its adaptation to cool climates and its high potential protein yield (Satovic et al. 2013).

Faba bean, like most other legumes, is symbiotic with nodule-forming bacteria with N2-fixing ability, which provides major benefits to cropping systems and the environment, and contributes to agricultural sustainability by soil improvement activities (Sprent 2009), so not only is it free of the need for nitrogen fertilizer, it also contributes to the nitrogen nutrition of the following crop. Faba bean is considered an excellent candidate to provide nitrogen input into temperate agricultural systems on account of its wide adaptation and high biomass productivity (Rispail et al. 2010).

Faba bean seeds are a good source of protein, fibre and nutritional value, and are widely grown for food and feed (Duc 1997). The protein content in faba bean is 24 – 35% of the seed dry matter and is very rich in lysine. It is one of the main sources of cheap protein for people in the Middle East, Latin America and Africa (consumed as dry or canned), and for livestock (mostly pigs and poultry) in many developed countries. Furthermore, the immature beans are widely used as a vegetable. However, the quality-limiting factors, such as vicine, convicin and tannins compromise the value of faba bean in many diets. Vicine and convicine are stored in the cotyledons and are harmful for humans carrying a genetic defect (Glucose-6-phosphate dehydrogenase deficiency, G6PD) and for both laying hen and broiler production, while tannins in the seed coat reduce protein and energy digestibility in monogastric animals (Crépon et al. 2010 and references in). Alleles conferring low tannin content (zt1 and zt2) and low vicine-convicine content (zvc) have been identified and are used in breeding programs (Crépon et al. 2010; Torres et al. 2010).

Though the agronomic and economic importance of faba bean is well documented, its cultivation is still limited due to its susceptibility to several biotic and abiotic stresses such as drought.

1.2 Drought adaptation

Dry lands cover over 40% of the world and are the home of around 2.1 billion poor people in nearly 100 countries (Mortimore et al. 2009; United Nations 2011). On a global basis, drought (temporary water deficit coupled with heat stress) is considered to be the most
important abiotic constraint to crop productivity (Fischer and Turner 1978; Huang et al. 2002). Climate change will lead to increased temperatures and changed precipitation patterns in the world. Models predict that climate change attributed to the emission of greenhouse gases will increase the frequency and intensity of droughts in agricultural areas worldwide (Oldfield 2005; Gornall et al. 2010; Dai et al. 2013). In order to improve crop performance, it is essential to understand how crop species respond and adapt to drought conditions. The key to drought adaptation for plant breeders and crop physiologists is tailoring the morphology, physiology and phenology of a crop to its environment in order to manage water economy. Two key aspects of drought physiology are regulation of water use through the stomata, and regulation of water uptake through the roots. This thesis focuses on leaf attributes associated with drought adaptation.

Faba bean, as an important source of protein in those dry areas of developing countries most likely to be impacted by climate change, is reputed to be relatively sensitive to water deficit among legume species (McDonald and Paulsen 1997; Khan et al. 2010). It is highly responsive to irrigation (Oweis et al. 2005), but in most production regions it is seldom if ever irrigated and mostly depends on stored soil moisture and rainfall for its growth and development. Thus, development of drought-adapted genotypes is an essential mission to improve yield and its stability in this crop, but progress has been relatively slow compared to other crops due to lack of efficient screening methods (Stoddard et al. 2006). Wide genotypic variation in water stress response has been reported in this species (Heringa et al. 1984; Grzesiak et al. 1997; Abdelmula et al. 1999; Amede et al. 1999; Link et al. 1999; Ricciardi et al. 2001; Khan et al. 2007), indicating the potential for breeding of drought-prone environments.

Field phenotyping for drought response is costly (when rain-exclusion shelters have to be established and maintained) and time-consuming, and often produces unreliable output due to seasonal variation, since periods of drought vary in length, timing and intensity (Khan et al. 2007). In contrast, screening under controlled conditions provides the potential for rapid response and uniform conditions. An ideal phenotyping method should be fast, low-cost, non-destructive, accurate and working on several samples (Richards et al. 2010; Monneveux et al. 2014). Then there should be a link to translate laboratory findings in controlled conditions to the field, and a novel genotype should also perform well in field conditions in order for the breeding program to be successful (Passioura 2012). Improvement in new sensor technologies has allowed the development of automated, high throughput and precise phenotyping platforms. Although costly to install, these platforms streamline and automate data collection in large numbers of samples and thus practically reduce manual logistics and some kinds of experimental errors. The LemnaTec (LemnaTec GmbH, Würselen, Germany; http://www.lemnatec.com/plant-phenotyping.php), PHENOPSIS (Granier et al. 2006), GROWSCREEN (Walter et al. 2006) and TraitMill (Reuzeau et al. 2006) are some of these platforms. Some platforms are specialised for certain types of plants or organs, and GlyPh, designed for high-throughput measurement of plant water use and growth in soya bean (Pereyra-Irujo et al. 2012), maybe adaptable for use in other legumes such as faba bean.

1.2.1 Response mechanisms

Three types of plant strategies for coping with drought stress have been characterized, namely drought avoidance, dehydration tolerance and escape (Ludlow and Muchow 1990;
Kramer and Boyer 1995). Each of these strategies has its own complex morphological and physiological mechanisms that involve complex genetic control.

*Drought avoidance* refers to mechanisms related to keeping high plant water status under water-limited conditions by minimizing water loss and/or maximizing water uptake through root characteristics. Stomatal closure is the most important drought avoidance mechanism and is affected by abscisic acid (ABA) transported from the roots (Schachtman and Goodger 2008). The guard cells are the structures and they open and close the stomata. Stomatal traits such as morphology (e.g., density and size) and function (e.g., conductance, \(g_s\)) are considered key determinants of plant growth and water status (reviewed in publication 1). Faba bean accessions with higher stomatal density were shown to have lower yield and less resistance to water stress, while lower stomatal density was associated with better drought avoidance (Ricciardi 1989). Appropriate stomatal activity might be the key role for improving drought adaptation in faba beans by reducing water loss and increasing transpiration efficiency (Dawish and Fahmy 1997). The \(g_s\) was found to be the most important trait determining water use in relatively small sets of faba bean accessions (Nerkar et al. 1981; Khan et al. 2007). Carbon isotope discrimination (\(\Delta^{13}C\)) has been proposed as a physiological criterion to select \(C_3\) crops, correlated with yield and water use efficiency, as a result of the balance between photosynthetic activity and \(g_s\) (e.g., Condon et al. 2004). Khan et al. (2007) found wide genetic variation for \(\Delta^{13}C\) in faba bean, and reported a negative correlation between transpiration efficiency and \(\Delta^{13}C\) under well watered conditions. Nevertheless, \(\Delta^{13}C\) is an expensive measurement, so it is suitable for important subsets of samples rather than large germplasm surveys. Leaf temperature has been shown to be a rapid and cost-efficient surrogate for measuring \(g_s\) and water status in several plant species (Blum 2011a) including faba bean (Khan et al. 2010). Water use efficiency (WUE) has often been proposed as one of the most important determinant of crop productivity under water-limited conditions and a target for plant drought adaptation. Blum (2009) argued, however, that selection for high WUE under drought conditions leads to dramatic losses in yield and drought adaptation. To overcome this, two layers of selection are needed: one for higher \(\Delta^{13}C\) and the other for shoot biomass (Krishnamurthy et al. 2013). Efficient use of water, i.e., maximizing soil moisture capture through transpiration, is the target for improvement of crop yield under drought conditions (Blum 2009).

Cuticular wax protects leaves against non-transpiration water loss. Water stress improved the deposition of cuticular wax and subsequently enhanced drought avoidance in transgenic alfalfa (*Medicago sativa* L.) (Zhang et al. 2005). In the same way, accumulation of cuticular waxes significantly increased in soya bean plants exposed to water deficit (Kim et al. 2007) and cuticular wax accumulation is considered an important drought avoidance strategy in many plants (Kosma et al. 2009; Zhou et al. 2013). However, abundant wax is not always correlated with drought tolerance (Ristic and Jenks 2002). Variation in the wax content, and in its synthesis in response to drought stress, have not yet been documented in faba bean.

Root characteristics have been shown to be an important contributor to drought avoidance in many crop species (Blum 2011a; Monneveux et al. 2014), including legume species such as groundnut (*Arachis hypogaea* L.), pigeon pea (*Cajanus cajan* (L.) Millsp.), cowpea (*Vigna unguiculata* (L.) Walp.), soya bean, common bean (*Phaseolus vulgaris* L.) (Purushothaman et al. 2013 and references in), chickpea (Kashiwagi et al. 2006; 2008a) and lentil (Sarker et al. 2005; Kumar et al. 2012; Singh et al. 2013). The uptake of water can be maximized and dehydration postponed by appropriate root characteristics. The first of these
is rooting depth, as deeper roots can continue to extract water from deep in the soil profile when shallow roots have already exhausted the water available to them (Blum 2011b). Faba bean is sometimes considered to be relatively shallow rooted, with maximum rooting depth of 50 – 90 cm, depending on soil type (Manschadi et al. 1998). Extensive variation in rooting depth exists in faba bean germplasm, for example, rooting depth of eight accessions at 35 days after sowing ranged from 32 to 45 cm (Khan and Stoddard, unpublished data; EU-Faba project). Previous work on faba bean, which has been focused on traits related to stomatal activity and other aspects of drought adaptation, has demonstrated the necessary need for research on roots in this species (Khan et al. 2007; 2010).

*Dehydration tolerance* refers to maintaining metabolic activities at low tissue water potential (surviving internal water deficits). Keeping high tissue turgor by osmotic adjustment (OA), antioxidant defence systems and changed dynamics of plant hormones are the major physiological stress-adaptive traits to minimize the detrimental effects of water deficit (Morgan 1984). OA maintains cell water content by increasing the osmotic force that can be exerted by cells on their surroundings and thus increasing water uptake (Blum 2011a). OA has not been deeply been investigated in faba bean, but it has been demonstrated in many other legumes (Morgan et al. 1991; Ashraf and Iram 2005; Turner et al. 2007; Silvente et al. 2012). In chickpea, OA was not related to changes in carbohydrate composition or leaf gas exchange under drought conditions (Basu et al. 2007). A wide variation was found in pea genotypes for OA when they were exposed to water stress, with some producing considerable amounts of soluble sugars and proline (Sánchez et al. 1998). Concerning faba bean, Amede et al. (1999) and Khan et al. (2007) reported a decline in osmotic potential in 19 and six inbred lines, respectively, under water stress conditions, but there was no evidence of osmotic adjustment. A wider spectrum of faba bean germplasm needs to be surveyed for the presence of osmotic adjustment.

*Drought escape* is the ability of a crop to complete its life cycle before the onset of drought or unfavourable conditions. The matching of crop growth to the pattern of rainfall is one of the most important breeding goals for drought-prone environments, particularly those prone to terminal drought. Early flowering is the most important phenological trait when terminal drought is likely. Consequently, early sowing along with early seedling vigour and better root establishment in dry environments allows rapid ground coverage and as a result reduces evaporation from the soil (Loss et al. 1997). Terminal drought escape via early production is a successful breeding strategy in Mediterranean environments (Araus et al. 2004 and references in). Although earliness allows crops to avoid terminal drought, it can lead to a yield penalty under well watered conditions (Monneveux et al. 2012). There is considerable variation in days to onset of flowering in faba bean accessions (Stoddard 1993; Khazaei et al. 2012) and developmental responses to temperature and day length (Iannucci et al. 2008; Patrick and Stoddard 2010). Earliness is also a desirable attribute for often short season climates such as boreal zone (Stoddard and Hämäläinen 2011).

At the end it should be noticed that combinations of these traits, not a single trait, are required to improve yield and its stability under drought conditions. For example, combined selection for earliness, increased harvest index and phenotypic plasticity has been proposed as a method to stabilize legume yield under Mediterranean conditions (Siddique et al. 2003).
1.3 Breeding

The situation of crop production in the 21st century has been analysed by many researchers and does not look bright (Pinstrup-Andersen et al. 1999). Furthermore many biotic and abiotic stresses significantly reduce crop yields. Plant breeding plays an important role for increasing crop yield and stress adaptation (Moose and Mumm 2008). Plant breeding describes methods for the creation, selection, and fixation of superior plant phenotypes for the development of improved cultivars. The primary goals of faba bean breeding are, those of any other crop, to improved yields, seed quality (e.g., low vicine-convicine and tannin-free), adaptation to biotic and abiotic stresses, and other agronomic traits (Torres et al. 2006; 2010; 2012). Relatively little effort has been made to breed faba bean for adaptation to abiotic stresses, especially using morpho-physiological traits related to drought adaptation.

1.3.1 Conventional breeding

Faba bean is partially cross-pollinated (allogamous) at frequencies ranging from 4-84%, with the value determined by the interaction between the plant genotype, its environment, and the population of pollinators (Bond and Poulsen 1983; Stoddard and Bond 1987). The mixed mating system means that breeding methods for self-pollinated or cross-pollinated crops are either not appropriate, or need adjustment. Two main breeding strategies are employed, population development with recurrent selection, and line breeding with management of the degree of cross pollination by isolation strategies (Gnanasambandam et al. 2012). Conventional faba bean breeding may take around 1 decade, from crossing inbred lines, population development, multi-environment evaluation of progenies for desired traits (including yield and its stability) and then selecting the superior genotypes. Faba bean can show high levels of heterosis (Bond 1989a; Link et al. 1996, 1997; Dieckmann and Link 2010; Meitzel et al. 2011). Part of this can be captured in a composite cultivar, where 2-3 inbreed lines are allowed to interbreed freely, and this was the main breeding strategy for several decades (Bond 1989a). Cytoplasmic male sterility (CMS) systems were developed, but suffered from instability, as an unacceptably high proportion of male-sterile individuals reverted to fertility (Bond 1989b; Link et al. 1997). ICARDA maintains two different categories of faba bean material, ILB (International Legume Bean) accessions from different countries, and BPL (Bean Pure Line) accessions that are derived through selfing from each chosen ILB line depending on the variation in it (Saxena and Varma 1985). There is potential to use homozygous BPLs as parents in faba bean breeding programmes to ensure the consistency of progeny.

Conventional breeding involves crossing whole genomes followed by selection of the superior recombinants from among the many segregation products. This procedure is laborious and time consuming, involving many crosses, several generations, and careful phenotyping, and linkage drag (tight linkage of undesired traits with the desired loci) may make it difficult to achieve the desired breeding outcome. DNA marker technologies, the development of several types of molecular markers and molecular breeding strategies offer possibilities to plant breeders and geneticists to overcome many of the problems faced during conventional breeding. Drought adaptation is a complex quantitative trait (Khan et al. 2010) controlled by several genes or QTLs of small effects, so many authors hold that breeding for it requires the combination of conventional and molecular strategies (Chaves et al. 2003; Blum 2011a).
The total area occupied for faba bean production worldwide dropped by 57% (5.4 to 2.3 M ha) from 1961 to 2012, but average yield has increased around 2-fold (0.9 to 1.7 t/ha) during the same time (Figure 2). The yield trend is still upward, showing how conventional breeding efforts have been successful in this crop (Bond et al. 1994).

Figure 2 World faba bean area harvested and yield from 1961 to 2012 (FAOSTAT 2014)

Faba bean yields are, however, considered to be unstable, which has been attributed to many causes, including the partial dependence on insect pollinators, and the general lack of improvement in breeding for resistance to numerous environmental constraints (biotic and abiotic stresses). Highly significant genotype × environment interaction for yield has been reported in faba bean accessions (Annicchiarico and Iannucci 2008; Flores et al. 2013). Breeding for adaptation to biotic and abiotic (particularly drought adaptation) stresses were summarized by Sillero et al. (2010) and Khan et al. (2010).

1.3.2 Genetic resources

There is considerable variation in plant genetic resources (PGR) collections conserved ex situ in agricultural genebanks, so screening genetic resources would be the first sound step of any pre-breeding program. Plant breeders find useful genes within these collections that can have impressive impacts on crop yields and adaptation to biotic and abiotic stresses. The efficient use of PGR is essential to conserve biodiversity, respond to future challenges including climate change, and thus pay big dividends to agriculture and environment (Johnson 2008).

A wealth of germplasm is available, comprising over 7.4 million accessions conserved ex situ in more than 1,700 national and international agricultural genebanks worldwide, and about 18% of the accessions are food legumes. Concerning faba bean, 43,695 accessions (accessed on 15 March 2013 from ICARDA genebank) are conserved mostly within 37 global genebanks. The ICARDA collection exceeds 9,000 accessions (21% of global
collection), followed by the Institute of Crop Germplasm Resources of the Chinese Academy of Agricultural Sciences, ICGR-CAAS (10%) and the Australian Temperate Field Crops Collection, ATFCC (6%) (FAO 2010). Faba bean is fifth among grain legumes for number of conserved ex situ accessions (after soya bean, ~230,000; chickpea, ~98,000; pea, ~94,000 and lentil, ~58,000 accessions).

Faba bean is represented in the collections by only the cultivated forms. Both botanical and molecular outcomes suggest that the wild ancestor of faba bean either has not yet been discovered or has vanished (Duc et al. 2010). Modernization of agriculture also reduces genetic diversity in crops, so ex situ collection saved in genebanks as well as in situ conservation by farmers in their agro-ecosystems are all the more crucial for present and future faba bean breeding programs. The next challenge concerning faba bean conservation in genebanks, is that due to the partial allogamy status of this species, maintenance activities are costly and difficult. Half of the keepers reported that their faba bean collections were maintained with open pollination (Suso et al. 2005).

Wide genetic variation has already been reported in faba bean genetic resources for agromorphological traits such as morphology (Robertson and El-Sherbeeny 1991; Perrino et al. 1991; Suso et al. 1993; Lang et al. 1993; Terzopoulos et al. 2003, 2004; Karaköy et al. 2014), seed size (Nadal et al. 2003), seed composition (Gresta et al. 2010; Khamassi et al. 2013; Karaköy et al. 2014), and yield and its components (Polignano et al. 1993; Keneni et al. 2005; Terzopoulos et al. 2008). Concerning abiotic stresses, high levels of adaptation have been reported for winter hardiness (Arbaoui et al. 2008a; Link et al. 2010) and drought (Amede et al. 1999; Khan et al. 2007). This enormous variability is expected in this ancient domesticated crop with its wide adaptability to various latitudes and altitudes, its ability to grow from spring and autumn sowings, and its various usages for food and feed.

In addition, molecular markers as a powerful and more reliable tool have been used for characterising genetic diversity in faba bean genetic resources. Different marker types have been used for this purpose such as RAPDs (Link et al. 2005). AFLPs (Zeid et al. 2001; 2003; Zong et al. 2009; 2010), TRAPs (Kwon et al. 2010), ISSRs (Terzopoulos and Bebeli 2008; Alghamdi et al. 2011; Wang et al. 2012) and SNPs (Kaur et al. 2014a). Exceptionally wide genetic diversity was found in accessions from North China, while those from Central China had less diversity and European and North African germplasm sets were genetically close to each other. This high level of genetic diversity, confirmed by both phenotyping and molecular markers, is a key to successful crop improvement.

Since agricultural genebanks conserve large numbers of accessions in overlapping collections, identifying suitable material is like searching for the proverbial needle in a haystack. Thus an efficient and rapid tool for mining genebanks is needed. Core collections were introduced as a way to work with fewer accessions that would represent most of the genetic variation of a large collection in a relatively small number (~5–10 % of the total collection) of chosen accessions (Frankel 1984). Nevertheless, core collections are not constructed to highlight or identify materials that are most likely to contain the specific adaptive traits desired in a particular search (see publication I and references in). In other words, the core collection concept focuses on genetic diversity, not adaptive phenotype.

The concept of centres of diversity that established the association between diversity and eco-geographic distribution was proposed in the early 1900s by Nicholas I. Vavilov (1887–1943), who recognized the importance of locating genetic variation based on eco-climatic
conditions (Vavilov 1922). Much earlier, in the 1600s, farmers in eastern North America were apparently aware of the need for ecological adaption and imported wheats from western Europe with better resistance to cold and pests (Olmstead and Rhode 2011). The Focused Identification of Germplasm Strategy (FIGS) was proposed by Mackay and Street (2004) as a new approach to utilize PGR. FIGS uses available environmental data (geographic and agro-climatic information) of collection sites to identify manageable sets of accessions “most likely” and “least likely” to express the trait of interest (Figure 3). This approach allows rapid reduction of the thousands of accessions to a manageable number with a high probability of carrying desired traits for the breeder’s strategy (Mackay and Street 2004; Publication I). Odong et al. (2013) classified core collections into three categories, namely those representing individuals (CC-I), extremes (CC-X) and distribution (CC-D). Among them, CC-X is similar to FIGS in the sense of considering geographical selection, but it still does not consider adaptive traits.

Figure 3 Distribution of core collection and FIGS sets in the overall spectrum of genetic diversity

Novel statistical approaches are becoming available to screen plant genetic resources more efficiently and precisely, such as machine learning algorithms (Bari et al. 2012; Bari et al. 2014a, b) and Bayesian mathematics (Dayanandan et al. 2013). The work of this thesis has run in conjunction with the testing of these methods for identifying potential sources of abiotic stress in faba bean (Bari et al. 2014c). At the same time, genotyping is becoming faster and cheaper. Next generation sequencing, NGS (i.e. high-throughput genotyping) is providing new tools to analyse PGR, such as identifying patterns of genetic diversity, mapping quantitative traits, and mining novel alleles that are not discovered yet (Kilian and Graner 2012). This would be a suitable approach for a species such as faba bean with its several breeding challenges.

1.4 Genomics

Faba bean is a diploid (2n = 12) with six remarkably large chromosomes, which makes it a suitable tool for cytogenetic studies. Its genome is one of the largest of any diploid field crops (~13,000 Mbp) and contains more than 95% of repetitive DNA. The faba bean genome size is 2.6, 3.2, 17.6 and 26 times larger than those of pea, lentil, chickpea and Medicago truncatula, respectively (Sato et al. 2010) and 4-fold larger than the human (Homo sapiens)
genome (Figure 4). Hence, certain kinds of investigation are less advanced in faba bean than in many other species. Sequence information is available on only 20,550 ESTs in the public databases, in contrast to nearly 1.5 million for soy bean (GenBank®, http://www.ncbi.nlm.nih.gov, 20.05.2014). Hence, functional genomic studies, identification of important genes and development of a highly saturated map are relatively limited in this species. Nevertheless, significant progress has been made to enrich its genetic and genomic resources in the last two decades. The lack of genome compatibility of faba bean with other Vicia species may explain why it is reproductively isolated from them. The mitochondrial genome sequence of faba bean (cv. Broad Windsor) was estimated at 588,000 bp. It comprised 52 genes, of which 32 encoded proteins, 3 rRNA, and 17 tRNA. Of the 114 unique open reading frames (ORFs) 36 lacked significant homology to known proteins (Negruk 2013).

Moreover, advances in NGS combined with techniques for the quantification of gene expression (e.g., SuperSAGE), are facilitating genome-wide transcriptome studies and target gene identification in organisms such as faba bean, from which massive nucleotide sequence information is not yet available. As a result, possible regulators revealed by SuperSAGE analysis and associated with the ascochyta-faba bean and ascochyta-lentil interactions have been genotyped and tentatively located on the crop genetic maps (Madrid et al. 2013) in order to identify potential targets for molecular breeding for ascochyta resistance. Furthermore, large-scale SSR marker development was successfully achieved using NGS on the faba bean genome (Yang et al. 2012).

**Figure 4** Genome size of faba bean compared to those of some other organisms
Table 1 History of molecular marker application for constructing genetic linkage map in faba bean

<table>
<thead>
<tr>
<th>Marker type / total marker</th>
<th>Mapping population / type</th>
<th>Population size</th>
<th>Total map length (cM)</th>
<th>Average inter-marker distance (cM)</th>
<th>No. of LGs (Assigned to chromosome)</th>
<th>Reference</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 RFLPs, 4 morpho, 3 iso, 3 RAPDs (17)</td>
<td>BC</td>
<td>–</td>
<td>231.4</td>
<td>–</td>
<td>7 (–)</td>
<td>van de Ven et al. (1991)</td>
<td>M</td>
</tr>
<tr>
<td>43 RAPDs, 7 iso, 1 RFLP (51)</td>
<td>2 F₁ families</td>
<td>20±44</td>
<td>300–350</td>
<td>–</td>
<td>11 (1)</td>
<td>Torres et al. (1993)</td>
<td>M</td>
</tr>
<tr>
<td>8 Morpho, 7 RFLPs, 4 iso, 4 RAPDs (23)</td>
<td>BC</td>
<td>–</td>
<td>400</td>
<td>–</td>
<td>7 (–)</td>
<td>Ramsay et al. (1995)</td>
<td>M, Q</td>
</tr>
<tr>
<td>147 RAPDs, 9 iso, 1 morpho (157)</td>
<td>7 F₂ families</td>
<td>813</td>
<td>850</td>
<td>–</td>
<td>48 (6)</td>
<td>Satovic et al. (1996)</td>
<td>M, T, C</td>
</tr>
<tr>
<td>105 RAPDs, 7 iso, 3 seed protein genes, 1 morpho (116)</td>
<td>3 F₂ families</td>
<td>175</td>
<td>~1,200</td>
<td>20</td>
<td>13 (7)</td>
<td>Vaz Patto et al. (1999)</td>
<td>M, T, Q, C</td>
</tr>
<tr>
<td>77 RAPDs (77)</td>
<td>F₁, RIL</td>
<td>57</td>
<td>973.2</td>
<td>14.66</td>
<td>14 (–)</td>
<td>Surahman (2001)</td>
<td>M</td>
</tr>
<tr>
<td>117 RAPDs, 2 iso, 2 seed protein genes (121)</td>
<td>F₂</td>
<td>196</td>
<td>1,445.5</td>
<td>13.72</td>
<td>16 (9)</td>
<td>Román et al. (2002, 2003)</td>
<td>M, Q</td>
</tr>
<tr>
<td>176 RAPDs, 6 iso, 4 SSRs, 3 seed protein genes, 2 morpho (191)</td>
<td>11 F₂ families</td>
<td>654</td>
<td>1,559</td>
<td>8</td>
<td>14 (5)</td>
<td>Román et al. (2004)</td>
<td>M, T, C</td>
</tr>
<tr>
<td>94 RAPDs, 4 iso, 3 SSRs, 2 seed protein genes (103)</td>
<td>F₂,3 families</td>
<td>159</td>
<td>1,308</td>
<td>–</td>
<td>18 (6)</td>
<td>Avila et al. (2004, 2005)</td>
<td>Q</td>
</tr>
<tr>
<td>135 ITAPS (135)</td>
<td>F₂, RIL</td>
<td>94</td>
<td>1,685.8</td>
<td>14.6</td>
<td>12 (–)</td>
<td>Elwood et al. (2008)</td>
<td>M, *</td>
</tr>
<tr>
<td>131 ITAPS (131)</td>
<td>F₂, RIL</td>
<td>101</td>
<td>1,635.39</td>
<td>14.73</td>
<td>21 (–)</td>
<td>Albano et al. (2008b)</td>
<td>M, Q</td>
</tr>
<tr>
<td>238 RAPDs, 21 Introm- spanning marker, 6 SSRs, 5 ESTs, 4 iso, 2 STSs, 1 SCAR (277)</td>
<td>F₂, RIL</td>
<td>165</td>
<td>2,856.7</td>
<td>12.72</td>
<td>21 (9)</td>
<td>Díaz-Ruiz et al. (2009 a,b; 2010)</td>
<td>Q</td>
</tr>
<tr>
<td>167 ESTs, 71 RAPDs, 11 SSRs, 3 RGAs, 3 seed protein genes, 2 iso, 1 morpho (258)</td>
<td>F₂, RIL</td>
<td>124</td>
<td>1,875</td>
<td>7.26</td>
<td>16 (8)</td>
<td>Cruz-Izquierdo et al. (2012)</td>
<td>M, Q, *</td>
</tr>
<tr>
<td>110 SSRs, 9 faba EST-SSRs, 5 grass pea EST-SSRs, 3 pea EST-SSRs, 1 ITAP (128)</td>
<td>F₂</td>
<td>129</td>
<td>1,587</td>
<td>12.5</td>
<td>15 (–)</td>
<td>Ma et al. (2013)</td>
<td>M</td>
</tr>
<tr>
<td>121 RAPDs, 38 ESTs, 6 SSRs, 5 RGAs, 1 defense-related gene, 1 seed protein gene (172)</td>
<td>F₁, RIL</td>
<td>119</td>
<td>1,402.1</td>
<td>9.87</td>
<td>29 (15)</td>
<td>Gutiérrez et al. (2013)</td>
<td>Q, C</td>
</tr>
<tr>
<td>729 markers in total</td>
<td>3 RILs (RIL1 Cruz-Izquierdo et al. (2012), RIL2 Díaz-Ruiz et al. (2009 a,b), RIL3 Gutiérrez et al. (2013))</td>
<td>408</td>
<td>4,613</td>
<td>6</td>
<td>43 (13)</td>
<td>Satovic et al. (2013)</td>
<td>M, Q, C</td>
</tr>
<tr>
<td>465 SNPs, 57 EST-derived SSRs (522)</td>
<td>F₁, RIL</td>
<td>95</td>
<td>1,216.8</td>
<td>2.3</td>
<td>12 (12)</td>
<td>Kaur et al. (2014b)</td>
<td>M, Q, *</td>
</tr>
</tbody>
</table>

LG linkage group, morpho morphological marker, iso isozymes, T Assignment of linkage groups to chromosomes by trisomic segregation, M Development of a map, Q QTL, C composite map, * comparative map
1.4.1 Marker assisted selection (MAS)

The use of DNA markers in plant breeding is called marker-assisted selection (MAS) and has opened a new gate in agricultural sciences, ‘Molecular breeding’, since the early 1980s. Use of new biotechnological methods holds promise to provide tools that assist plant breeders to develop new cultivars with greater reliability and cost-efficiency. There are three major types of genetic markers, namely morphological, biochemical and DNA markers. Three main advantages of selection based on DNA markers over conventional phenotyped-based selection are the relative simplicity of MAS that may save time and resources, the potential to use it at earlier stages of plant growth than phenotyping, and the lack of effect of environment (Collard et al. 2005; Collard and Mackill 2008). MAS is based on the selection of molecular markers in a segregating population and/or pyramiding desirable alleles at the target loci. This approach was first used at faba bean breeding in the early 1990s and has had numerous applications since, such as construction of linkage maps (Table 1), assessing the level of genetic diversity within germplasm (see Alghamdi et al. 2012; Kaur et al. 2014a and references in), quantitative trait loci (QTL) mapping, and detection of economically important genes (Table 2). The current constraints of using MAS in faba bean breeding are the early stages of its DNA marker technology development, the reliability and accuracy of QTL mapping studies, and most importantly, the limited range of markers and limited polymorphism of markers in breeding material, in contrast to some other species.

Sirks (1931) was the earliest to report on the faba bean genome, identifying 19 genetic factors that formed four linkage groups. His material was lost during World War II. Four decades later, Sjödin (1971) used translocation lines in the assignment of different loci (morphological observations) to their respective chromosomes. Several morphological markers and seven isozyme loci were subsequently added into the faba bean map (Suso et al. 1993; Torres et al. 1995 and references in). Primary trisomics were also used to assign genes and linkage groups to specific chromosomes (Torres et al. 1998; Van Patto et al. 1999). The first DNA-base linkage map in this species was constructed with just 17 markers, of which 10 were DNA-based (van de Ven et al. 1991). The faba bean genetic linkage map was initially developed using morphological, isozyme, RFLP, RAPD and seed protein gene markers that were limited in number, transferability or synteny with model plant species. The first SSR markers were described in this species by Požárková et al. (2002) and soon enriched a composite map (Román et al. 2004). The first gene-base map was reported in 2008 (Elwood et al. 2008) and more recently enriched by using 511 genic markers (Kaur et al. 2014b) with 2.3 cM average inter-marker density (Table 1). Satovic et al. (2013) published a reference consensus genetic map that covered 4,062 cM in six main linkage groups corresponding to the six chromosomes of faba bean.

1.4.2 Markers based on DNA

DNA markers linked to desired traits are a valuable selection tool in modern plant breeding. Molecular markers have been sought for both simple and complex traits (reviewed in Torres et al. 2010; 2012). In the case of major genes, the bulked segregant analysis (BSA) method has been used for the identification of markers that are tightly linked to the genes underlying parameters such as tannin content (controlled by zt1 and zt2, two independent but complementary genes), vicine-convicine content (Gutiérrez et al. 2006; 2007; 2008), hypersensitive response to rust (Avila et al. 2003), and indeterminate growth habit (Vf-TFL1) (Avila et al. 2006; 2007). The performance of these markers is still not completely
satisfactory because of their genetic distance from the trait and their lack of conservation in all genetic backgrounds. Compared with BSA, the use of a candidate gene approach allows the development of diagnostic or perfect markers that are completely linked to or within the selected trait. This approach is possible only when the target gene is well characterized in the target species or in a related species carrying orthologous genes.

Genic molecular markers are developed from the coding sequences such as ESTs or fully characterized genes, frequently with known functions, in contrast to most of the previously developed molecular markers (e.g., RAPDs and AFLPs) that are directly from the total genomic DNA without any information on function (Varshney et al. 2007). The efficacy of genic molecular markers may thus be superior to DNA-based markers in molecular breeding.

Gene-based SNPs are the most abundant class of polymorphic sites in any organism’s genome, providing low genotyping cost per data point, locus specificity (accuracy and reproducibility), co-dominance, simple documentation and potential for high-throughput analysis. Thus they have become a powerful tool in genetic mapping, association studies, assessing genetic diversity and positional cloning (Rafalski 2010; Semagn et al. 2014). The discovery of SNPs in candidate genes or transcript sequences (ESTs) has been a recurrent strategy in plant genomics, mainly because SNPs on which the markers are based could themselves be causative for trait differences (Galeano et al. 2012). SNPs were recently used to assess genetic diversity in faba bean germplasm (Cottage et al. 2012a; Kaur et al. 2014a). Elwood et al. (2008) and Kaur et al. (2014b) have developed exclusively gene-based genetic linkage maps in faba bean. Less progress has been made to develop and characterize EST-derived SSRs in this species by screening the NCBI database (Gong et al. 2010; Ma et al. 2011; 2013; Akash and Myers 2012) than has been achieved in the other cool-season food legumes, lentil (e.g., Kaur et al. 2011; Sharpe et al. 2013), chickpea (e.g., Garg et al. 2011; Hiremath et al. 2011; Jhanwar et al. 2012; Stephens et al. 2014) and pea (e.g., Franssen et al. 2011; Kaur et al. 2012).

1.5 Synteny

Functional genomics in faba bean is more poorly developed than in many other field crops, especially soya bean, common bean and chickpea among legumes, largely because of its large genome. An indirect approach such as synteny with a model species may overcome this shortcoming. Comprehensive comparative genomics to assess synteny can assist back-and-forth use of genomic resources between different legume species, and also help to decrease cost and increase the efficiency in genetics and crop breeding by speeding up gene discovery (Ellwood et al. 2008). Comparative genomic analysis has established the correspondence between faba bean chromosomes and those of the model legumes (Ellwood et al. 2008; Cruz-Izquierdo et al. 2012).

Arabidopsis thaliana as a well known model species has played an important role in illuminating the function of many genes, especially in monocotyledons (Meinke et al. 1998). Nevertheless, the genomic information of A. thaliana may not be easily translated for use in legumes, since it is susceptible to different pathogens, is incapable of biological nitrogen fixation, and lacks many of the secondary pathways and quality traits that are widespread in legumes (Handberg and Stouggaard 1992). Hence, M. truncatula and Lotus japonicus have been developed as model legumes (Sato et al. 2007; 2010), as both are diploid, with haploid chromosome numbers of 8 and 6, respectively. The genomes are less than 4% of the size of
the faba bean genome (500 Mbp and 472 Mbp, respectively) (Sato et al. 2010). In addition, they are small self-pollinating plants with short life cycles and able to be transformed by such methods as using Agrobacterium tumefaciens or A. rhizogenes, making them ideal model legume species (Rispail et al. 2010). Nearly 264 and 315.1 Mb of M. truncatula (http://www.medicagohapmap.org/?genome) and L. japonicus (http://www.kazusa.or.jp/lotus/) have been sequenced and there are more than 260,000 and 240,000 ESTs available in the public domain (Benson et al. 2013, GenBank®, http://www.ncbi.nlm.nih.gov, last visit 15.05.2014).

Considerable synteny among legume genomes has been reported (e.g., Choi et al. 2004; Kalo et al. 2004; Zhu et al. 2005; Phan et al. 2006; 2007; Elwood et al. 2008; Kalo et al. 2011; Cruz-Izquierdo et al. 2012). The recent soya bean (Schmutz et al. 2010) and chickpea (Varshney et al. 2013) whole genome sequences are also being exploited in comparative mapping studies, in order to identify orthologous genes and further develop efficient molecular tools in other species such as faba bean.

1.5.1 Medicago truncatula

Comparative genomics studies have shown that M. truncatula (also known as: barrel medic, barrel medick or barrel clover) has high collinearity with faba bean (Ellwood et al. 2008; Cruz-Izquierdo et al. 2012; Kaur et al. 2012; 2014b). The high level of conservation in gene order indicates that M. truncatula can be used to facilitate faba bean molecular breeding (Rispail et al. 2010). An NCBI BLAST search showed that about 45% of faba bean mitochondrial DNA sequence was homologous to M. truncatula nuclear sequences (Negruk 2013). The high synteny between M. truncatula and faba bean can be used in orthogonal gene identification, MAS, and constructing highly saturated genetic linkage maps or physical maps. Identification of QTLs and then candidate genes also may be used to produce transgenic elite lines for specific breeding purpose by genetic engineering.

From use of synteny between M. truncatula and pea, 5,460 pea unigenes were identified through comparative mapping and placed in the consensus map of pea (Bordat et al. 2011). A high syntenic resolution map between M. truncatula and faba bean using 768 SNPs is in progress to construct a consensus genetic map in faba bean (Webb et al. in preparation).

Both faba bean and M. truncatula are cool-season legumes. M. truncatula is considered to be relatively drought-adapted, compared to other legumes (González et al. 1995) and pea (Gálvez et al. 2005). The genetic similarity and synteny between M. truncatula and faba bean may help in identifying genes responsible for drought adaptation in faba bean and accelerate breeding for drought-prone environments. ESTs of genes involved in drought adaptation in chickpea, pea and M. truncatula have been already identified (Jayashree et al. 2005; Buhariwalla et al. 2005).

1.6 Quantitative trait loci

Many agriculturally important traits (e.g., yield, quality factors and many forms of disease resistance) are controlled by many genes and are called quantitative traits. Detecting the major loci of these traits, QTLs, is one of the most important aspects of MAS application in plant breeding. In QTL mapping, it is generally assumed that there is a small number of major genes detectable in relatively small samples. Combining phenotypic data and allelic marker segregation along a genome allows the detection of QTLs, which increases our
knowledge on their inheritance and gene action. The summary of detected QTLs in faba bean is presented in Table 2.

The first QTL mapping in this species was reported by Ramsay et al. (1995), who detected several loci for morphological and biochemical traits. Mapping QTL in faba bean for biotic stresses, such as resistance to parasitic plants or pathogenic fungi, has been attempted (Table 2). Two of the major disease constraints, namely ascochyta blight (caused by Ascochyta fabae Speg.,) and crenate broomrape, have been widely subjected to QTL studies on F2 and RIL populations. The QTLs accounting for significant proportions of these resistances have been identified, mapped and validated in multi-environment trials (Roman et al. 2003; Avila et al. 2004; Díaz-Ruiz et al. 2009a, b; 2010; Gutiérrez et al. 2013; Kaur et al. 2014b). In addition, some efforts has been given to rust resistance (Avila et al. 2003), but relatively little for chocolate spot (Botrytis fabae Sard.,) in spite of the importance and widespread nature of this disease. QTLs for these traits have been assigned to 5 of the 6 chromosomes (I, II, III, V and VI), but no linkage group had been assigned to chromosome IV until recently, when Ruiz-Rodríguez et al. (2014) assigned linkage group I.B (previously reported by Cruz-Izquierdo et al. 2012) to this chromosome.

A little progress has also been made in identifying QTLs for abiotic stresses such as frost tolerance (Arbaoui et al. 2008b), yield components, and early flowering time associated with drought avoidance during the reproductive stage (Vaz Patto et al. 1999; Avila et al. 2005; Cruz-Izquierdo et al. 2012), but outcomes still are far behind those gained in biotic stresses. Saturation of the genomic regions associated with target regions and QTL validation in multiple environments and genetic backgrounds are prerequisites to uncover reliable marker-trait associations. Nevertheless, marker density in faba bean remains rather low for the identification of responsible QTLs / genes and use of MAS approaches. Considerably fewer QTLs have been reported in this species than in other grain legumes (Kumar et al. 2011). At present, linkage groups detected in faba bean (mostly by AM Torres’s research group, see Ruiz-Rodríguez et al. 2014) have been anchored to all faba bean chromosomes, which provides the source for integration of genetic and physical maps that will facilitate fine QTL mapping and gene identification in this species.
<table>
<thead>
<tr>
<th>Trait</th>
<th>No. of QTL/loci</th>
<th>Chromosome</th>
<th>Main output</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Morphology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphological trait</td>
<td>–</td>
<td>unknown</td>
<td>The first QTL map in faba bean including morphological and biochemical traits</td>
<td>Ramsay et al. (1995)</td>
</tr>
<tr>
<td>Growth habit</td>
<td>1</td>
<td>V</td>
<td>The first molecular marker (CAPs) for growth habit (ii) by using candidate gene approach (F&lt;sub&gt;2&lt;/sub&gt;)</td>
<td>Avila et al. (2006)</td>
</tr>
<tr>
<td>Seed coat colour</td>
<td>3c</td>
<td>II</td>
<td>Complete gene sequence of Vf-TFL1 using different genotypes</td>
<td>Avila et al. (2007)</td>
</tr>
<tr>
<td><strong>Biotic stress</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broomrape resistance <em>Orobanche crenata</em> Forsk.</td>
<td>3</td>
<td>I, II, VI</td>
<td>QTLs Oc1, Oc2, Oc3 explained more than 70% of phenotypic variation</td>
<td>Roman et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>I, II, VI</td>
<td>QTLs Oc2, Oc3, Oc4, Oc5 in 2 years and 3 environments</td>
<td>Diaz-Ruiz et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>VI</td>
<td>2 new QTLs added, Oc7 and 8</td>
<td>Gutierrez et al. (2013)</td>
</tr>
<tr>
<td><em>Orobanche foetida</em> Poir.</td>
<td>2</td>
<td>I, III</td>
<td>Mapping Of1 and Of2</td>
<td>Diaz-Ruiz et al. (2008b)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>V</td>
<td>Co-localization of Oc8 and Of3 in chromosome V confirms a common resistance against <em>O. crenata</em> and <em>O. foetida</em>.</td>
<td>Gutierrez et al. (2013)</td>
</tr>
<tr>
<td>Ascochyta fabae* Speg.</td>
<td>2</td>
<td>II, III</td>
<td>QTLs associated to ascochyta blight resistance, Af1 and Af2 explained ~ 50% of phenotypic variation</td>
<td>Roman et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>III</td>
<td>Isolate and organ-specific QTLs for ascochyta blight resistance, Af3 to Af8</td>
<td>Avila et al. (2004)</td>
</tr>
<tr>
<td>Uromyces viciae-fabae* (Pers.) J. Shört.</td>
<td>1</td>
<td>unknown</td>
<td>Identifying the first molecular markers linked to <em>Uvf-1</em> gene conferring hypersensitive resistance against rust (F&lt;sub&gt;2&lt;/sub&gt;)</td>
<td>Avila et al. (2003)</td>
</tr>
<tr>
<td><strong>Abiotic stress</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frost tolerance</td>
<td>5</td>
<td>unknown</td>
<td>Frost tolerance and physiologically related traits</td>
<td>Arbaoui et al. (2008b)</td>
</tr>
<tr>
<td>Yield and its components</td>
<td>7</td>
<td>VI</td>
<td>First putative QTLs for seed weight, explained 30% of phenotypic variation</td>
<td>Vaz Patto et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>unknown</td>
<td>Mapping floral characters, yield distribution and yield components</td>
<td>Avila et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>II, V, VI</td>
<td>QTLs for days to flowering, flowering length, pod length, No. seeds and ovules per pod</td>
<td>Cruz-Izquierdo et al. (2012)</td>
</tr>
<tr>
<td><strong>Quality traits</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low vicine-convicine <em>vc</em></td>
<td></td>
<td>unknown</td>
<td>RAPD markers linked to low vicine-convicine content and their conversion to SCARs (F&lt;sub&gt;2&lt;/sub&gt;)</td>
<td>Gutierrez et al. (2006)</td>
</tr>
<tr>
<td>Zero tannin</td>
<td>zt1</td>
<td>unknown</td>
<td>RAPD markers linked to <em>zt1</em> and their conversion to SCARs (F&lt;sub&gt;2&lt;/sub&gt;)</td>
<td>Gutierrez et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>zt2</td>
<td>unknown</td>
<td>RAPD markers linked to <em>zt2</em> and their conversion to SCARs (F&lt;sub&gt;2&lt;/sub&gt;)</td>
<td>Gutierrez et al. (2008)</td>
</tr>
</tbody>
</table>
2 AIMS OF THIS STUDY

The main aim of this study was to investigate the genetic potential for drought adaptation in faba bean by developing tools for germplasm utilization and marker assistant selection (MAS) that can be used in pre-breeding. To reach this end, two collections of faba bean, each of 201 accessions, originating from the wet and dry regions of the world and chosen using a novel approach (FIGS), were screened for leaf traits related to drought adaptation (stomatal morphology and function, and leaf water relations). Advanced biparental populations were developed to detect QTLs governing these traits, and synteny with *M. truncatula* as a model legume species was employed to identify candidate genes. During population development, the genetic control of some Mendelian characters (e.g., stipule spot pigmentation, SSP) was also considered.

The specific aims of this study were:

1) to compare the drought adaption-related traits of two faba bean sets originating from environments with contrasting seasonal moisture availabilities, selected using machine-learning algorithms. The primary hypothesis was that ecotypic differences occurred so traits associated with plant moisture regulation and lifecycle would differ between the two sets (Publication I).

2) to test whether faba bean germplasm from drought-prone and non-drought-prone environments differed in drought adaption related traits under both well watered and water deficit conditions. In addition, inter-relationships between studied traits were investigated (Publication II).

3) *a*, to develop a gene-based linkage map using SNPs; *b*, to identify QTLs controlling drought adaption-related traits; and *c*, to exploit synteny between faba bean and *M. truncatula* to examine the putative gene content of the detected QTLs for candidate genes (Publication III).

4) to investigate the genetic control of SSP and to exploit synteny between faba bean and *M. truncatula* to detect putative candidate genes (Publication IV).
3 MATERIALS AND METHODS

The experimental part of the work is described here as a general outline.

Table 3 Summary of methods used in this work are listed below. More details are described in the original publications (I – IV) and following text

<table>
<thead>
<tr>
<th>Method</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused Identification of Germplasm Strategy (FIGS)</td>
<td>I, II</td>
</tr>
<tr>
<td>Machine-learning algorithms*</td>
<td>I</td>
</tr>
<tr>
<td>Screening drought adaptation-related phenotypes</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Evaluation of water deficit response</td>
<td>II, III</td>
</tr>
<tr>
<td>Multivariate statistics</td>
<td>II</td>
</tr>
<tr>
<td>SNP genotyping</td>
<td>III, IV</td>
</tr>
<tr>
<td>DNA extraction*</td>
<td>III, IV</td>
</tr>
<tr>
<td>QTL analysis</td>
<td>III</td>
</tr>
<tr>
<td>Synteny analyses with <em>M. truncatula</em></td>
<td>III, IV</td>
</tr>
<tr>
<td>Developing a genetic linkage map</td>
<td>III</td>
</tr>
<tr>
<td>Crossing and population development</td>
<td>III, IV</td>
</tr>
<tr>
<td>KASPar assay*</td>
<td>III, IV</td>
</tr>
<tr>
<td>Backcrossing</td>
<td>IV</td>
</tr>
<tr>
<td>Microsatellite loci screening (confirmation of hybrids)</td>
<td>IV</td>
</tr>
</tbody>
</table>

Methods marked with asterisk were not carried out by HK
3.1 Germplasm survey

3.1.1 Construction of the FIGS sets

The “dry” and “wet” sets each comprised 200 accessions (listed in supplementary Table 1 publications I and II), plus one of known response from earlier investigations (ILB 938/2 and Aurora/2, respectively) (Khan et al. 2007; Khazaei et al. 2010). Collection sites were identified where the long-term average annual rainfall was between 300 and 550 mm/year (dry set) and over 800 mm/year (wet set), respectively. After the sites were identified, one accession per site was chosen at random. A hierarchical cluster analysis was performed using the collection site agro-climatic parameters (long-term yearly precipitation, long-term yearly aridity index, long-term yearly minimum temperature, long-term yearly maximum temperature, temperature seasonality, precipitation seasonality, precipitation of wettest quarter and precipitation of coldest quarter; http://www.worldclim.org/bioclim) in order to achieve a balanced representation of sites (more details can be found in publication I and II).

3.1.2 Germplasm screening

The two sets were evaluated using randomized complete block design with 4 replicates for some morphological, physiological and phenological traits during 2010–2011 under well watered conditions (measurements listed in Table 4). In the next step, on the basis of principal component analysis (PCA) on studied traits, ten representative accessions from each set (wet and dry, called “subsets”) were selected for exposure to transient moisture stress following Khan et al. (2007) (publication II).

3.2 Molecular breeding

3.2.1 Mapping population development and selection

Four parental lines [Aurora/2 (drought-sensitive), Mélodie/2 (High EUW), Disco/2 (zt2) and ILB 938/2 (high WUE)] differing in water-deficit response and other valuable traits were chosen at the start of the project for crossing and development of RIL populations. Crosses Mélodie/2 × ILB 938/2, Mélodie/2 × Disco/2, ILB 938/2 × Aurora/2 and ILB 938/2 × Disco/2 were prepared by hand, along with their reciprocals, in the insect-proof greenhouse. The standard sets of quantitative-genetics generations (F1, F1, F2, and BC1F1) were prepared. The F2 seeds were advanced to the next generation (F3–F4) by single seed descent, each comprising ~350 individuals. The parental lines were in a collection of 35 accessions used in the development of 756 SNPs that have been validated in an international panel (Cottage et al. 2012b; Webb et al. in preparation). Mélodie/2 and ILB 938/2 showed the highest genetic distance from each other, with 222 polymorphic SNPs, so this cross was advanced to F3 and chosen for deeper investigation (publication III and IV). Mélodie/2 was selected from a French low vicine-convicine cultivar with high EUW, while ILB 938/2 was a selection from an Ecuadorian landrace with relatively high WUE.

3.2.2 Genotyping

High throughput genomic DNA was extracted by LGC Genomics laboratory (LGC Genomics, Hoddesdon, UK) according to the manufacturer’s instructions.
SNP genotyping was accomplished using the KASPar™ (Kompetitive Allele Specific PCR) assay (KBioscience, UK) platform, a single-plex SNP genotyping methodology using allele-specific amplification followed by fluorescence detection for genotyping (http://www.kbioscience.co.uk/reagents/KASP_manual.pdf). The SNP sequences and specific alleles for parental lines used in this study are listed in Supplementary Table 2, publication III.

3.2.3 Phenotyping

The F₅ population (Mélodie/2 × ILB 938/2) and its with parental lines were phenotyped in a climate-controlled glasshouse during 2013–2014 for some morphological and physiological traits (see Table 4) using a completely randomized design with 3 replicates under well watered and water deficit conditions (publication III). Furthermore, the population was phenotyped for vicine and convicine content (data not shown in this thesis).

3.2.4 QTL mapping, synteny and candidate genes

Details of genetic linkage map construction and QTL analysis are presented in Table 5. SNP segregation was subjected to the chi square goodness-of-fit test to assess deviations from the expected Mendelian segregation ratio (1:1). SNPs showing normal diploid segregation ($P \geq 0.05$) were chosen for the linkage map construction. The Kosambi function was used to calculate the map distance in centiMorgans (cM) (Kosambi 1943). Composite interval mapping (CIM) was used to detect the relationship between each linkage group and putative QTL locations. Significant QTLs were analysed with CIM. Cofactors were determined using the forward and backward method in the standard CIM model with the probability in and out of 0.1. The 95% significance threshold for QTL detection was determined by the phenotype permutation using 1,000 permutations at an experiment-wise $P < 0.05$.

The overall level of collinearity between the detected putative QTLs (using sequences flanking SNPs) and M. truncatula was examined. Reciprocal best BLAST hit E-values were employed to determine the strength of the orthologous relationship. Gene family, phylogenetic context, and tissue-specific gene expression profiles of the orthologous genes from model plant species were examined using the LegumeIP database (http://plantgrn.noble.org/LegumeIP/) (M. truncatula, Gene Model, Mt3.5v3).

3.3 Measurements

Details of measurements are listed in Table 4. For the leaf morphology (stomatal morphology and leaflet area), gas exchange traits and leaf temperatures the youngest, fully expanded leaflets of 8-week-old plants were used in all experiment (publications I, II and III).

3.3.1 Induction of water deficit (Publications II and III)

For publication II, transient drought response (subsets in 2011), 2 L plastic pots were filled with 1.42 kg of the mixture (sand / peat, 3:1 v/v) that had a water holding capacity (WHC) of 20% (w/w), using a completely randomized factorial design with 4 replicates. Each pot was brought to WHC by adding 285 ml of water. Pots were weighed every 2 days and amounts of water equal to the loss in weight were added. Ten days after sowing, 60 g of perlite was
added to the top of each pot (to reduce soil evaporation). In the well watered treatment, the pots were irrigated as described above until harvesting at 13 weeks (when the main tiller of the plants turned yellow). Half of the plants were exposed to a gradual and uniform water stress starting at 5 weeks by reducing water by 2% (w/w) of available water per every two days to bring the moisture level down from the field capacity (20% w/w) to moisture stress (2–4% w/w). Pots were weighed and where water use exceeded 2%, irrigation was applied. Measurements were done 2 weeks after the gradual drying was started.

For publication III in 2013, 4 L pots were filled with 2.48 kg of the mixture (sand / peat, 2:1 v/v) that had a WHC of 24% (w/w). Each pot was brought to WHC by adding 595 ml of water. Pots were weighed three times per week and amounts of water equal to the loss in weight were added to maintain soil moisture level at field capacity. Immediately after 8-week-old plants were phenotyped under well watered conditions, a uniform water stress was started by reducing relative available water by 2% (w/w) every day to bring the moisture level down from the field capacity (24% w/w). Pots were weighed every day and where water use exceeded 2%, irrigation was applied (Khan et al. 2007). When the soil moisture content had reached 2–4%, 10 days after the induction of water deficit was started, canopy temperature was measured. Thereafter normal irrigation was continued until harvest.

3.4 Growth conditions and statistical analysis

All experiments (germplasm survey and mapping population phenotyping) were conducted in the climate-controlled glasshouse of the Department of Agricultural Sciences, University of Helsinki, Finland. Soil moisture level was maintained at field capacity with automatic irrigation for all plants in the well watered experiments. Seeds of all accessions were inoculated with *Rhizobium leguminosarum* biovar. *viciae* (faba bean strain, Elomestari Oy, Tornio, Finland) before sowing. A mixture of sand (size: 0.5–1.2 mm, Saint-Gobain Weber Oy Ab, Helsinki, Finland) and peat (White 420 W, Kekkilä Oy, Vantaa, Finland) containing all essential nutrients was used as the potting medium. For all experiments, at three and five weeks after sowing, 70 ml of nitrogen-free fertilizer (equivalent to 20 kg of P and 24 kg of K per ha) was added to each pot. Photoperiod was adjusted to 14 h light and 10 h dark, and the temperature was maintained at ≈ 21°C day/15°C night ±2°C. Photosynthetic photon flux density (PPFD) was approximately 300 µmol m⁻² s⁻¹ at the canopy level. Relative humidity was maintained at 60%. The temperature, humidity and light conditions were automatically recorded throughout the experiments. For the germplasm survey experiments, 2 L pots used with mixture of sand and peat (3:1 v/v) was used. In the population phenotyping experiment, 4 L plastic pots filled by sand and peat (2:1 v/v) used. Statistical analysis and software used to analyse data in this thesis are presented in Table 5.
Table 4 Measurements and methods used in the experiments

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Germplasm survey</th>
<th>Mapping population</th>
<th>Instrument / Method / Reference</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stomatal morphology</td>
<td>Yes (SD, SL, SW, SA)</td>
<td>Yes (SD, SL)</td>
<td>Impression method (Wang and Clarke 1993), using abaxial surface of leaflet</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Leaflet area</td>
<td>Yes</td>
<td>No</td>
<td>LI-6200 (LI-COR Inc., Lincoln, NE, USA)</td>
<td>I, II</td>
</tr>
<tr>
<td>Number of tillers</td>
<td>Yes</td>
<td>No</td>
<td>Counting</td>
<td>I</td>
</tr>
<tr>
<td>Gas exchange traits</td>
<td>Yes (A_{net}, g_s, E, WUE)</td>
<td>Yes (g_s)</td>
<td>LI-6400 (LI-COR Inc.), equipped with a 2x3 cm leaf chamber with a LED light source (6400-02B, 90% red and 10% blue), reference CO₂ concentration = 400 µmol mol⁻¹, flow rate = 400 µmol s⁻¹, PPFD = 1,000 µmol m⁻² s⁻¹</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Leaf temperature*</td>
<td>Yes</td>
<td>No</td>
<td>LI-6400 (LI-COR Inc.)</td>
<td>I, II</td>
</tr>
<tr>
<td>Canopy temperature*</td>
<td>Yes</td>
<td>Yes</td>
<td>Infrared thermometer (IRT, FLUKE® thermometer gun 574, Everett, WA, USA)</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Water use subsets</td>
<td>No</td>
<td>No</td>
<td>Weighing pots (according to Khan et al. 2007)</td>
<td>II</td>
</tr>
<tr>
<td>Biomass WUE subsets</td>
<td>No</td>
<td>No</td>
<td>Total biomass / water used</td>
<td>II</td>
</tr>
<tr>
<td>Relative water content</td>
<td>Yes</td>
<td>No</td>
<td>According to Barrs and Weatherley (1962)</td>
<td>I, II</td>
</tr>
<tr>
<td>Days to flowering</td>
<td>Yes</td>
<td>No</td>
<td>Number of days to the onset of flowering</td>
<td>II</td>
</tr>
<tr>
<td>Seed weight</td>
<td>No</td>
<td>Yes</td>
<td>Weighing 10 mature dried seeds per plant</td>
<td>III</td>
</tr>
<tr>
<td>Stipule spot pigmentation</td>
<td>No</td>
<td>Yes</td>
<td>Scoring the absence / presence of pigmentation</td>
<td>IV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Germplasm survey</th>
<th>Mapping population</th>
<th>Instrument / Method / Reference</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>unit area of leaflet, A_{net}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>photosynthetic rate, g_s stomatal conductance, E transpiration rate, WUE_i Intrinsic water use efficiency (A_{net}/g_s), PPFD photosynthesis photon flux density</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* In publications I and II, the canopy temperature differences from air temperature and leaflet temperature differences from air temperature were calculated
<table>
<thead>
<tr>
<th>Statistics / Function</th>
<th>Software</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>One way ANOVA</td>
<td>R</td>
<td>II</td>
</tr>
<tr>
<td>Two way ANOVA</td>
<td>R and SPSS v. 21</td>
<td>II, III</td>
</tr>
<tr>
<td>Machine-learning algorithm</td>
<td>R [caret, randomForest, svm (e1071) and ksvm (kernalab libraries)]</td>
<td>I</td>
</tr>
<tr>
<td>Correlation and Regression</td>
<td>R</td>
<td>II, III</td>
</tr>
<tr>
<td>T-test</td>
<td>R</td>
<td>I</td>
</tr>
<tr>
<td>Contrast analysis</td>
<td>SPSS v. 21</td>
<td>II</td>
</tr>
<tr>
<td>PCA</td>
<td>R</td>
<td>II</td>
</tr>
<tr>
<td>Skewness, kurtosis and Shapiro-Wilk test</td>
<td>R</td>
<td>III</td>
</tr>
<tr>
<td>Chi-square</td>
<td>R</td>
<td>III, IV</td>
</tr>
<tr>
<td>Linkage map construction</td>
<td>MapDisto v. 1.7.7.8. (LOD = 3, recombination fraction = 0.3) (Lorieux 2012)</td>
<td>III, IV</td>
</tr>
<tr>
<td>Composite interval mapping</td>
<td>Windows &quot;QTL Cartographer v 2.5_011&quot; (control marker = 5, window size = 10, walk speed = 1) (Wang et al. 2012)</td>
<td>III, IV</td>
</tr>
<tr>
<td>Linkage groups and QTL drawing</td>
<td>MapChart v. 2.2 (Voorrips 2002)</td>
<td>III, IV</td>
</tr>
</tbody>
</table>

*a* Analysis of variance  
*b* R Development Core Team (2012)  
*c* IBM® SPSS® Statistics for Windows, v. 21, Armonk, NY, USA  
*d* Models used in this study: Classification and regression training, random forest and support vector machines (publication I)
4 RESULTS AND DISCUSSION

4.1 Germplasm screening

The morpho-physiological and phenological measurements of the dry and wet FIGS sets differed significantly under well watered conditions. The dry set had significantly larger stomata, greater stomatal area per unit area of leaflet, higher relative water content (RWC), leaflet area and transpiration rate, and in contrast lower leaflet and canopy temperature, fewer fertile tillers and fewer days to flowering than the wet set. Greater variation for most traits was observed within the wet set. Stomatal density and stomatal area per unit area of leaflet were negatively correlated with gas exchange parameters and positively correlated with intrinsic water use efficiency, while stomatal size showed the opposite trends (publications I and II).

Three machine-learning algorithms (classification and regression training, random forest and support vector machines) were employed to examine difference between the two FIGS sets (publication I). The outcome revealed that sets were classified into two groups based on the phenotypic evaluation with little or no overlap between sets, showing a high degree of discrimination between sets. Among the traits, leaflet and canopy temperature and RWC contributed most to the discrimination between sets (Table 5 in publication I). Machine-learning algorithms have been little used in this context, so their results were confirmed by using a well known multivariate statistical methods, principal component analysis, that further assessed the pattern of variation in the two sets and confirmed the results of the machine-learning algorithms results. The first four principal components (PC1-4) explained 81.3% of the total variation, of which the first component (PC1) accounted for around half. Gas exchange traits (photosynthetic rate, gs, intrinsic WUE and transpiration rate) made the highest contribution to PC1. The second PC was highly dependent on leaflet and canopy temperature along with RWC. These traits contributed more than other measurements to discriminate the sets from each other (Figure 5), and also made the greatest contribution to separating the two sets according to machine-learning algorithms (see publication I). The next components were mostly based on stomatal morphology. Distribution of the two FIGS sets shows how FIGS successfully demonstrates that ecotypic differentiation driven by moisture availability occurred in faba bean germplasm (Figure 5). Similarly, several salt-tolerant accessions in wild Phaseolus species were identified from arid, coastal, or saline areas close to their centre of diversity in Central and South America (Bayuelo-Jiménez et al. 2002). When applied to Lathyrus species, FIGS also allowed selection of manageable subsets with a high probability of finding accessions adapted to drought and heat (Shehadeh 2011). Furthermore, FIGS has been used to detect numerous sources of resistance to biotic stresses in wheat (Triticum ssp.) and barley (Hordeum vulgare L.) (e.g., Kaur et al. 2008; Bhullar et al. 2009; Endresen et al. 2011; Bari et al. 2012, 2014b).
Figure 5 Principal component analysis scatter diagram based on morpho-physiological traits in the wet set (black circles) and dry set (grey circles) of faba bean germplasm.

Faba bean has been grown for hundreds or thousands of years in a wide range of environments, including those subjected to often severe drought constraints (Duc et al. 2010), so we expected notable variation between sets from contrasting seasonal moisture availabilities. Ecosystem disturbance has been shown to add further selection pressure in annual plant evolution (Peleg et al. 2005). Other methods demonstrated the existence of ecological adaptation in several plant species (De Micco and Aronne 2012; Sarris and Koutsias 2014). All of these results support the hypothesis that accessions with desirable adaptive traits can be identified through the process of predicting the environmental profiles that select for these traits, and the question was whether FIGS would adequately predict these profiles, and would do so better than earlier methods. Analysis of natural genetic variation leads to the discovery of novel genes and alleles, especially in the field of plant adaptation to the environment (Shindo et al. 2007). Crop wild relatives are the most promising sources of novel alleles to improve drought adaptation in crops (Nevo et al. 2002), but in the case of faba bean, there is no wild relative available, so novel methods for utilizing PGR such as FIGS play a key task in this important aspect. In contrast, the use of core collections often misses adaptive and rare alleles (Gepts 2006; Pessoa-Filho et al. 2010).

Evaluation of drought adaptation requires phenotyping by exposing plants to contrasting water regimes. So, the next step sought natural variation within and between the wet and dry sets for response to water deficit using manageable subsets, 10 accessions representing the average of each set based on PCA. A moderate, transient, mid-season water deficit was used, as this is typical of what a crop may experience in the field almost anywhere in the world (Araus et al. 2002). Water deficit affected most traits and a significant set × water stress interaction was observed (publication II). Most traits, including canopy temperature depression (Figure 6), clearly separated two sets under well watered conditions, and the
response of most traits to water deficit also clearly differed between the two sets. The most outstanding drought adaptation capacity was found in accessions from the dry set, directing future exploration to these ecosystems. Genetic variation has been shown in leaf temperature RWC, Δ¹³C and g, of a small set of drought-adapted and drought-sensitive lines of faba bean (Khan et al. 2007). Our result showed that water stress caused stomatal densities to increase significantly in the dry set while stomatal size decreased in both sets. Stomatal area and density were negatively correlated to each other in all experiments, so the stomatal area per unit of leaflet remained fairly constant in the two sets. Stomata do not move. The results relates partly to reduced leaf size in both sets during the drought interval, due to reduced cell elongation from low turgour. Differences in stomatal density between the sets suggests that something else occurred, either more differentiation of guard cells in the dry set, or greater noise in one set that obscured the differences. These findings were in agreement with those on several other species (Franks and Farquhar 2007).

Under water deficit conditions, gas exchange fell strongly. The clearest differences between sets were in leaf temperatures and RWC. The dry set accessions had cooler canopies under well watered conditions and their leaf temperatures increased more under water deficit than those of the wet set (Figure 6). This demonstrates that accessions from dry regions have more effective stomatal control than those from wet regions in both wet and dry conditions. Overall, these results confirm the key role of stomatal function in water relations in this species. In the field, lower canopy temperature under drought stress conditions indicates a better capacity to take up water from deep soil and hence better water status (Blum 2011a). Fast and accurate phenotyping criteria are currently seen as a major issue for improvement of drought adaptation in crops (Richards et al. 2010). Canopy temperature can be measured by an infrared thermometer (IRT) as a rapid and cost-efficient alternative for preliminary screening for drought adaption under controlled and uniform conditions in faba bean (Khan et al. 2007, 2010). Canopy temperature has also been shown as an efficient field phenotyping tool for drought adaptation in other legume species (e.g., soybean, McKinney et al. 1998; chickpea, Kashiwagi et al. 2008b; cowpea, Hall 2012). Canopy temperature measured by IRT as a screening technique for dehydration avoidance was first reported by Blum et al. (1982). Under drought conditions, this is the most drought adaptive trait associated with yield performance and has high heritability (Oliwares-Villegas et al. 2007). Good practice in using the IRT, such as those protocol at http://www.plantstress.com/methods/IRT_protocol.htm, should be followed carefully to avoid large error variance and non-repeatable results. Water stress caused a reduction in RWC in both sets, but it was greater reduction in the wet set (27%) than in the dry set (19%). The ability of genotypes to maintain RWC when moisture is limiting is a good indicator of drought adaptation (Barrs and Weatherley 1962; Anyia and Herzog 2004; Blum 2011a).

In the current study we screened a large number of faba bean accessions (402), which showed a wider range of variation for morpho-physiological measurements than previous reports, suggesting a high potential for breeding for these traits. For example, in our germplasm g, ranged from ~ 0.040 to 0.620 mol m⁻² s⁻¹, in contrast to the range of 0.125 to 0.176 mol m⁻² s⁻¹ in six accessions known to differ for drought response (Khan et al. 2007). Similarly, ten faba bean accessions (Tanzarella et al. 1984) showed a much narrower range (77% less) for stomatal density than our work. These comparisons clearly show the enhanced identification of genotypic variation that results from the large germplasm set analysed in this work, and from using FIGS to seek extremes in a trait.
**Figure 6** Effects of water treatments on the canopy temperature depression in dry and wet set: well watered (*white column*); water deficit (*gray column*). Error bars show mean ± 1 SEM

### 4.2 Molecular breeding

#### 4.2.1 Response of parental lines to water deficit

The parents used for map construction showed different responses to water deficit. Mélodie/2 had a cooler canopy under well watered conditions and a sizeably higher increase in canopy temperature under water deficit conditions than ILB 938/2. The g₄ showed the reverse trend. Water deficit had 3-fold more effect on biomass production in Mélodie/2 than in ILB 938/2 but biomass in Mélodie/2 under water deficit conditions was the same as ILB 938/2 under well watered conditions (publication III). These results show the ability of ILB 938/2 to maintain higher water status under water deficit conditions, as it was selected for high WUE with relatively low yield. In contrast Mélodie/2 with higher EUW showed better productivity under drought conditions than ILB 938/2, with better water uptake ability through a well developed root system (Khazaei and Stoddard unpublished data). These two parental lines nicely explain the concept of WUE and EUW in sense of Blum (2009). Selection of parental lines, which are genetically and geographically far enough from each other, is the key point to build a promising RIL populations for successful genetic mapping and QTL detection (Würschum 2012).

#### 4.2.2 Linkage map and chromosome assignment

Two hundred and eleven individuals were randomly chosen (of 350) from the population Mélodie/2 × ILB 938/2 for phenotyping and genotyping (publication III). After considering
rejected markers based on the grounds of excess missing data (> 5% missing data per marker) or excessive expected chi-squared goodness of fit ratio, 188 SNPs, as along with seed coat colour (publication III), hilum colour and SSP (publication IV) were successfully mapped in nine linkage groups that covered ~ 928 cM of faba bean genome. The linkage groups were composed of 2–43 loci and their length ranged from 3.07 to 223.02 cM, with an average inter-marker density of 5.8 cM. The map showed a high degree of synteny with the genome of *M. truncatula* (publication III). High-throughput SNPs have greatly accelerated the development of high density linkage maps, with 2.3 cM average marker density having been achieved in faba bean recently (Kaur et al. 2014b; more information in Table 1).

The number of genetic linkage groups should be the same as chromosome number in the species, so six LGs is expected in faba bean. Nevertheless, the number of LGs generally exceeds the chromosome number, ranging from 7 to 48 in faba bean, and map length varied from 231.4 to 4,613 cM (Table 1). Our genetic map had seven main linkage groups and two fragments which is slightly more than the appropriate number, and the map was slightly smaller than some others (Table 1). These differences may be related to factors such as the large size of the genome, the number and types of markers, the software analysis used for construction of the genetic map (e.g., recombinant fraction and LOD value), the number and generation of progeny, and especially the genetic background and genetic distance of the parents (Young 2000; Broman 2006; Semagn et al. 2006; Cheema and Dicks 2009).

The correspondence of linkage groups to their respective chromosomes was performed using markers with identified chromosome location or linkage according to Cruz-Izquierdo et al. (2012) and the current consensus genetic map published by Satovic et al. (2013). The assignments made here are consistent with those possible using the higher syntenic resolution possible using a 768-locus consensus genetic map (Webb et al. in preparation) (Figure 7). Seven of the nine linkage groups were assigned to the six faba bean chromosomes (I–VI) (publication III).
Figure 7 Synteny between M. truncatula (Mt) and faba bean (Vf) chromosomes. This figure was extracted from Webb et al. (in preparation)

4.2.3 Putative QTLs

There is an urgent need to identify chromosome regions associated with the climate change-related traits in crops in order to assure food security in the near future. Progress in MAS for a biotic stresses such as drought in faba bean is relatively modest compared to that on biotic stresses, although some progress have been made for yield and its components (Table 2 and Torres et al. 2010). There is no corresponding previous report on morpho-physiological traits related to drought adaptation such as stomatal morphology and function, so the QTLs detected in this study would be a good starting point for future studies. QTLs were identified on two of the six faba bean chromosomes, II and IV. Fourteen QTLs were detected from five studied traits (stomatal density and length, gs, canopy temperature and seed weight) under well watered conditions over two years (2013 and 2014) and one QTL for canopy temperature under drought conditions in 2013. The LOD threshold at $P = 0.05$ varied from 2.8 to 3.0. The proportion of phenotypic variation explained by each individual QTL ranged from 5.7% to 9.3%. Thirteen QTLs were located on chromosome II, with eight of them clustered in the same genomic region ($qSD-2013-1$, $qSD-2014$, $qSL-2013$, $qSL-2014$, $qgs-2013$, $qCT-2012$, $qCT-2013$ and $qCT-2014$), in the marker interval, Vf_Mt4g014430_001. These traits also showed strong phenotypic correlations.

All traits were phenotyped in two years under well watered conditions, and the detected QTLs remained stable between years except that for $g_s$. Canopy temperature was also
phenotyped under water deficit conditions in 2013 and its QTL (qCTd-2013) was mapped in the same chromosome, chromosome II, ~ 125 cM from those found under well watered conditions (publication III). The direction of additive effect illustrates that alleles driving warmer leaves under well watered conditions probably came from Mélodie/2 and those conferring cooler leaves under stress conditions were from ILB938/2, which is in agreement with the concept of WUE and EUW (Blum 2009).

A QTL was detected in chromosome II for seed weight, which remained stable during the two years of phenotyping (qSW-2013-1 and qSW-2014-1). The second stable QTL governing seed weight was located on chromosome IV (qSW-2014-1 and qSW-2014-2). QTLs for this trait have been previously reported in chromosome VI (Vaz Patto et al. 1999). This difference is at least partially attributable to the different parents used, with the study by Vaz Patto et al. (1999) using paucijuga and major lines whereas the present study used equina (ILB 938/2) and minor (Mélodie/2) lines, and partly to the different distribution of markers in the two studies.

Many QTLs governing flowering time and reproductive traits associated with drought avoidance in faba bean were already mapped on chromosomes I, IV, V and VI (Cruz-Izquierdo et al. 2012; Satovic et al. 2013; Table 1). In this work, however, transient rather than terminal drought was the objective, and it was a glasshouse-based experiments, assessment of grain yield was not considered appropriate. Yield is related to high gs, photosynthetic rate and cooler canopies in several crops (Fischer et al. 1998; Khan et al. 2010), so the region in chromosome II that made a significant contribution to stomatal function warrants further investigations.

A short-term mapping population (e.g., F2 and BCs) developed through the BSA approach is considered a good starting point for molecular mapping, while a long-term mapping population such as one comprised advanced RIL is suitable for characterization of traits of importance. Several QTLs in this species were detected at the F2 generation, especially before 2008, and some of them were later validated at advanced generations, e.g. ascochyta blight resistance (Díaz-Ruiz et al. 2009a) (see Tables 1 and 2). In more recent studies, mapping populations are generally advanced to the F5 or F7 generation (Table 1 and 2). In this study, we used an F3 RIL population and Takuno et al. (2012) pointed out that QTL mapping does not necessarily require RILs at F5 or higher generations, and that F4 (or even F3) populations would be almost as useful as highly advanced populations.

Two QTLs governing stomatal density and gs (qSD-2013-2 and qgs-2014) were mapped in regions of distorted SNPs (publication III). The precision of the linkage map (map distances and marker ordering) may be biased in regions of many distorted markers (Lorieux et al. 1995a, b). Additionally, QTL regions often overlap with regions undergoing segregation distortion, because favourable QTL alleles are in excess frequencies in those regions (Boopathi 2013). Therefore, QTLs found in these regions are more uncertain than those located elsewhere.

4.2.4 Synteny

Despite the large difference between the genome size of faba bean and M. truncatula, a tight relationship was found between the two genomes (Figure 7), with strong evidence for extensive collinearly between linkage groups as it suggested by Rispail et al. (2010). Based on annotated genic SNPs derived from M. truncatula, candidate genes were found within
detected QTL intervals. Candidate genes for *M. truncatula* and Arabidopsis were almost identical, and were mostly related to growth and development, stress and transport (publication III). A segment of *M. truncatula* chromosome IV that harbours receptor-like protein kinase (RLKs) was consistently linked with most of the detected QTLs, including those for stomatal density and length, g, and canopy temperature under well watered conditions. The candidate gene for canopy temperature under water deficit was ribose-phosphate pyrophosphokinase 4 in *M. truncatula* chromosome VI. RLKs are the best known stress-induced genes and control a wide range of physiological responses such as growth regulation and specific adaptations in several plant species (e.g., Casson and Gray 2008; Marshall et al. 2012; Osakabe et al. 2013). Gene ontology for both intervals with detected QTLs for seed weight were related to regulation of seed germination (Arginyl-tRNA-protein transferase 1) (Manahan and App 1973; Holman et al. 2009) and embryonic development ending in seed dormancy (Splicing factor 3B subunit 4) (Raab and Hoth 2007; Aghamirzaie et al. 2013; http://www.medicagohapmap.org/?genome).

There is another success story for candidate gene discovery in faba bean. An ortholog of CEN/TFL1-like genes was recognized to be responsible for determinate growth habit in faba bean, and this was confirmed in several plant species and later used to develop a CAPs marker for use in MAS (Avila et al. 2007).

### 4.3 Mendelian traits

During population development, we noticed that parental lines used for population development also differed in several Mendelian traits, such as seed coat, hilum and flower colour, and stipule spot pigmentation (SSP). These characters were recorded during population development (F1 to F3) as well as in backcrosses, their heritability was investigated and they were mapped along with SNPs. The most detailed findings were on genetic control of SSP by using four segregating populations in the F1 to F5 generations (Mélodie/2 × ILB 938/2, Mélodie/2 × Disco/2, ILB 938/2 × Aurora/2 and ILB 938/2 × Disco/2) as well as reciprocal crosses. The results revealed that two unlinked genes, *sspl* and *ssp2*, determined this trait. A novel locus, *sspl*, decoupling pigmentation in the flower from that in the extra-floral nectarines in ILB 938/2 was detected and mapped onto chromosome I along with hilum colour using the Mélodie/2 × ILB 938/2 population (publication IV). Another gene, *ssp2* (z2) in Disco/2 was already mapped in this species (Gutierrez et al. 2008). In addition, the SSP as an early-expressed morphological marker confirmed hybridity in the crosses, so this trait can be recommended for use as a simple marker in programs focused on low tannin or vicine-convicine along with white flower and colourless hilum. Since *sspl* is recessive, but rare (it was present in 6% of the 402 faba bean accessions), it could be used in the long run if a sufficient number of parents carry it. Another practical application might be to use the *sspl* gene as a co-marker for visual evidence of outcrossing during pedigreed seed production if a particular cultivar carries *sspl* as a homozygote. If the same cultivar were homozygous recessive for a trait that was much more difficult to track, for example, low vicine-convicine, the appearance of nectary spots would be a simple co-marker in many situations. If this were desirable, a closely linked molecular marker could be used to design crosses appropriately for pyramiding because heterozygotes and homozygotes would be distinguishable. The *sspl* locus was tightly linked into a well conserved region of *M. truncatula* chromosome VI with some candidate Myb and bHLH (basic helix-loop-helix) transcription factor families that have been shown to regulate the anthocyanin pigmentation.
pathway in many species (e.g., Quattrocchio et al. 1998; Tai et al. 2013). Green seed coat colour was mapped on faba bean chromosome IV (publication III), while black seed coat colour was shown to be on faba bean chromosome II (Sjödin 1971; Díaz-Ruiz et al. 2010; Cruz-Izquierdo et al. 2012).
5 CONCLUSION AND FUTURE PROSPECTS

This study has combined the novel use of FIGS for identifying sources of abiotic stress resources and synteny with *M. truncatula* for gene mapping and discovery in faba bean. An appropriate phenotyping method, mapping population and suitable marker system are the key requirement for a successful molecular breeding program. The differences between the two FIGS sets (drought-adapted and non-drought-adapted) allowed demonstration of a reliable and cost-efficient phenotyping tool for screening traits related to drought adaptation (*i.e.* canopy temperatures). In the next step, using a mapping population with wide variation at both DNA sequence and phenotypic level and using *M. truncatula*-derived SNPs, all brought success in detecting the genetic regions and even candidate genes (*e.g.*, RLKs) for studied traits with potential for further functional genomics analysis. RLKs have already been widely shown to be related to abiotic stress responses in model plants as well as in several crop species. These results confirmed that genetic information from model species can be easily translated to faba bean as a “genome orphan”.

When traits other than yield associated with drought adaptation are considered for mapping, poor phenotyping can often cause failures (Varshney et al. 2011) and hence generate confusing results. A fast and reliable phenotyping method is needed for the particular crop in the specific drought profile. Significant correlations allow utilizing simpler traits as indirect selection criteria for yield and drought adaptation. In this species, leaf temperatures seems a reliable phenotyping method but it is highly recommended that a subset of samples be subjected to a more accurate screening method such as Δν13C. The most severe droughts mostly occur in poor countries with little access to expensive equipment such as sensor technologies or even Δν13C, so economical tools like IRT are exceptionally important.

Our findings support the assertion that FIGS can be an effective tool in pre-breeding programmes for traits related to drought adaption in faba bean, and this may be extruded to other traits related to abiotic stresses and adaptation to climate change. The core-collection approach asks "what is the range of variation in the species" while the FIGS approach asks "where can we find the extremes of variation in this trait in the species". Thus FIGS is complementary to core collections. The major demand for germplasm collections from agricultural genebanks is mostly for adaptive traits (Kenneth Street personal communication), so in contrast to core collections, FIGS represents a dynamic, direct and practical approach that focuses on the sought-after trait rather than on generalized measures of genetic diversity. Overall, CGIAR suggests the FIGS approach, alongside current core collection activities, should be part of breeding strategies. Several FIGS sets developed by ICARDA are available to speed up research on climate change adaptive traits (http://www.icarda.org/figs-%E2%80%93-new-approach-mining-agricultural-genebanks-speed-pace-innovation-and-food-security).

This work was carried out under climate-controlled greenhouse conditions and these findings should be confirmed under field conditions. When preparing to introduce detected putative QTLs as MAS tools, they should be validated in different environments. The identification of candidate genes takes away some of the apparent complexity of drought adaptation in faba bean. Co-localisation of these genes with known QTLs involved in drought adaptation may help select regulatory genes involved in this respect and to define a rational approach to developing drought-adapted cultivars. Candidate genes also can be
transferred to elite lines of faba bean by MAS or genetic transformation. Limited progress has been shown on gene transfer or pyramiding using MAS in legumes, but some progress has been made in soya bean, pea and common bean (Dwivedi et al. 2007).

Research on the root, the hidden part of plant, lags behind that on shoots. Physiologists and breeders are becoming positioned to breed crops with better rooting systems that improve productivity under water-limited conditions. Preliminary results on root phenotyping showed that drought-adapted accessions had relatively better rooting systems (branching, density and architect) than the wet adapted set, but root length did not differ. More research needs to be conducted on root morphology and function and its relation to shoot characteristics in this species, since drought adaptation in rice, for example, was associated with a more extensive root system as well as reduced stomatal density and improved WUE (Yu et al. 2013).
ACKNOWLEDGMENTS

I am very lucky to have Dr. Frederick L. Stoddard as my main supervisor. His profound knowledge in crop science, breeding, genetics and statistics along with his kind support, guidance and encouragements culminated this work. I am very grateful to my co-supervisor Professor Mikko J. Sillanpää for his advices and supports during molecular biology work. Many thanks to Professor Pirjo Mäkelä leader of crop science group, for her support for my PhD course management, thesis review and pre-reviewer suggestions and also Dr. Arja Santanen for her advice and practical help during lab work. My PhD monitoring team, Professor Teemu Teeri and Dr. Mervi Seppänen are acknowledged for their valuable comments and suggestions. I would like to thank Dr. Petr Smykal and Dr. Simo Hovinen, pre-reviewers of the thesis for their constructive comments.

Many interesting ideas to develop the current work came from our international collaboration for developing the FIGS sets by Dr. Kenneth Street, Dr. Abdallah Bari (ICARDA) and Dr. Michael Mackay (The University of Queensland, Australia), and for developing the SNPs platform by Professor Donal M. O’Sullivan (University of Reading, UK). I wish to thank Professor Wolfgang Link (Georg-August-University, Germany) for providing the seeds of the parental lines for population development.

I owe great gratitude to Markku Tykkyläinen, Sanna M. Peltola, Sini Lindström and Daniel Richterich, technical staff of the glasshouses of the University of Helsinki, for their kind assistance during the experiments.

I would like also to acknowledge Damian Wach, Guillermo Mínguez Vélaz, Sandro Maria Cavaleri, and Stefano Zanotto, visiting students, and Yao Lu, MSc student of this university, for their for their kind assistance at different stages of this project. Many thanks to all friends and colleagues at the crop science group, HY.

This work was financially supported by CIMO (the Centre for International Mobility), the Emil Aaltonen Foundation (Emil Aaltosen Säätiö), Niemi Säätiö, Department of Agricultural Sciences, University of Helsinki and the Doctoral Program in Plant Science. Part of the research reported here was conducted in conjunction with "Legume Futures (Legume-supported cropping systems for Europe)", a collaborative research project funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No. 245216.

I give special thanks to my wife, Aida for being patient and kind during the work and thesis writing, and to my parents for their encouragements during my studies which led me to the doctoral degree.

Helsinki, September 2014

Hamid Khazaei
6 REFERENCES


Bari A, Khazaei H, Street K, Sillanpää MJ, Endresen DTF, De Pauw E, Chaubey YP, Dayanandan S, Stoddard FL, Damania AD (2014c) *In silico* evaluation of plant genetic resources to search for traits for adaptation to climate change. (Submitted manuscript)


Blum A (2009) Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. Field Crops Res 112:119–123

Blum A (2011a) Plant breeding for water limited environments. Springer-Verlag, New York. 252 p


Bond DA (1989a) Prospects for commercialisation of F1 hybrid field beans *Vicia faba* L. Euphytica 41:81–86
Bond DA (1989b) A short review of research on male sterility and prospects for F₁ hybrid varieties in field beans (Vicia faba L.). Euphytica 41:87–90


Boopathi NM (2013) Genetic mapping and marker assisted selection: Basics, practice and benefits. Springer India


FAO (2010) The second report on the state of the world’s plant genetic resources for food and agriculture. Rome, Italy. 398 p


Raina SN, Rees H (1983) DNA variation between and within chromosome complements in Vicia species. Heredity 51:335–346


Stoddard FL, Hämmäläinen K (2011) Towards the world’s earliest maturing faba beans. Grain Legumes 56:9–10


Vavilov NI (1922) The law of homologous series in variation. J Genet 12:47–89


